Accuracy of 1-Hour Plasma Glucose During the Oral Glucose Tolerance Test in Diagnosis of Type 2 Diabetes in Adults: A Meta-analysis

Diabetes Care 2021;44:1062–1069 | https://doi.org/10.2337/dc20-1688

Vasudha Ahuja,[1] Pasi Aronen,[2] T.A. Pramodkumar,[3] Helen Looker,[4] Angela Chetrit,[5] Aini H. Bloigu,[6] Auni Juutilainen,[7] Cristina Bianchi,[8] Lucia La Sala,[9] Ranjit Mohan Anjana,[3] Rajendra Pradeepa,[3] Ulagamadesan Venkatesan,[3] Sarvanan Jebarani,[3] Viswanathan Baskar,[3] Teresa Vanessa Fiorentino,[10] Patrick Timpel,[11] Ralph A. DeFronzo,[12] Antonio Ceriello,[9] Stefano Del Prato,[8] Muhammad Abdul-Ghani,[12] Sirkka Keinänen-Kiukaanniemi,[6,13] Rachel Dankner,[5,14] Peter H. Bennett,[4] William C. Knowler,[4] Peter Schwarz,[11,15,16] Giorgio Sesti,[17] Rie Oka,[18] Viswanathan Mohan,[3] Leif Groop,[1,19] Jaakko Tuomilehto,[20,21,22] Samuli Ripatti,[1,23,24] Michael Bergman,[25] and Tiinamaija Tuomi[1,19,26]

OBJECTIVE

One-hour plasma glucose (1-h PG) during the oral glucose tolerance test (OGTT) is an accurate predictor of type 2 diabetes. We performed a meta-analysis to determine the optimum cutoff of 1-h PG for detection of type 2 diabetes using 2-h PG as the gold standard.

RESEARCH DESIGN AND METHODS

We included 15 studies with 35,551 participants from multiple ethnic groups (53.8% Caucasian) and 2,705 newly detected cases of diabetes based on 2-h PG during OGTT. We excluded cases identified only by elevated fasting plasma glucose and/or HbA_{1c}. We determined the optimal 1-h PG threshold and its accuracy at this cutoff for detection of diabetes (2-h PG ≥11.1 mmol/L) using a mixed linear effects regression model with different weights to sensitivity/specificity (2/3, 1/2, and 1/3).

RESULTS

Three cutoffs of 1-h PG, at 10.6 mmol/L, 11.6 mmol/L, and 12.5 mmol/L, had sensitivities of 0.95, 0.92, and 0.87 and specificities of 0.86, 0.91, and 0.94 at weights 2/3, 1/2, and 1/3, respectively. The cutoff of 11.6 mmol/L (95% CI 10.6, 12.6) had a sensitivity of 0.92 (0.87, 0.95), specificity of 0.91 (0.88, 0.93), area under the curve 0.939 (95% confidence region for sensitivity at a given specificity: 0.904, 0.946), and a positive predictive value of 45%.

CONCLUSIONS

The 1-h PG of ≥11.6 mmol/L during OGTT has a good sensitivity and specificity for detecting type 2 diabetes. Prescreening with a diabetes-specific risk calculator to identify high-risk individuals is suggested to decrease the proportion of false-positive cases. Studies including other ethnic groups and assessing complication risk are warranted.

META-ANALYSIS

In 1979, the National Diabetes Data Group and the World Health Organization (WHO) established the current practice of diagnosing type 2 diabetes based on fasting or 2-h threshold levels after a 75-g oral glucose tolerance test (OGTT) (1,2). The diagnostic criteria have since undergone two major changes by the WHO and American Diabetes Association (ADA): first, lowering of the diagnostic threshold of the fasting plasma

[1]*Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland*
[2]*Biostatistics Unit, Faculty of Medicine, University of Helsinki and Helsinki University Hospital, Helsinki, Finland*
[3]*Madras Diabetes Research Foundation & Dr. Mohan's Diabetes Specialties Centre, ICMR Centre for Advanced Research on Diabetes and IDF Centre of Excellence in Diabetes, Chennai, India*
[4]*Phoenix Epidemiology and Clinical Research Branch, National Institute for Diabetes and Digestive and Kidney Diseases, Phoenix, AZ*
[5]*Unit for Cardiovascular Epidemiology, Gertner Institute for Epidemiology and Health Policy Research, Ramat Gan, Israel*
[6]*Center for Life Course Health Research, University of Oulu, Oulu, Finland*
[7]*University of Eastern Finland, Kuopio University Hospital, Kuopio, Finland*
[8]*Section of Diabetes and Metabolic Diseases, Department of Clinical and Experimental Medicine, University Hospital of Pisa, Pisa, Italy*
[9]*Department of Cardiovascular and Dysmetabolic Diseases, IRCCS MultiMedica, Milan, Italy*
[10]*Department of Medical and Surgical Sciences, University Magna Graecia of Catanzaro, Catanzaro, Italy*
[11]*Department of Medicine III, Technical University of Dresden, Dresden, Germany*
[12]*Division of Diabetes, University of Texas Health Science Center at San Antonio, San Antonio, TX*
[13]*Healthcare and Social Services of Selänne, Pyhäjärvi, Finland*
[14]*Department of Epidemiology and Preventive Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel*
[15]*Paul Langerhans Institute of the Helmholtz Zentrum München at the University Hospital Carl Gustav Carus and the Medical Faculty of TU Dresden (PLID), Dresden, Germany*

glucose (FPG) from 7.8 mmol/L to 7.0 mmol/L in the late 1990s, and second, introduction of HbA_{1c} as an additional diagnostic criterion in the late 2000s (3–6).

A similar consensus does not exist for diagnosing prediabetes, also referred to as intermediate hyperglycemia (IH). WHO, ADA, and an ad hoc "International Expert Committee" advocate different criteria to define IH based on FPG (impaired fasting glucose [IFG]), 2-h plasma glucose (2-h PG) during the OGTT (impaired glucose tolerance [IGT]), and/or HbA_{1c} (7). Nevertheless, multiple studies in various ethnicities have indicated that 1-h plasma glucose (1-h PG) ≥8.6 mmol/L is a more accurate predictor of incident type 2 diabetes than IFG, IGT, HbA_{1c}, or their combination (7). Hence, an expert panel proposed a 1-h PG ≥8.6 mmol/L level to define IH (7). Since several studies have shown the association of 1-h PG with cardiovascular disease and mortality and a better and independent association of postchallenge glucose concentration than FPG and HbA_{1c} with these outcomes, it is logical to evaluate the potential of 1-h PG for the detection of type 2 diabetes (8–12).

Zhou et al. (13) and Paddock et al. (14) reported 1-h PG threshold values for detection of type 2 diabetes in Chinese and American Indian populations, respectively. As the threshold could be affected by study design, differences in recruitment of participants, and ethnicity as well other factors, we performed a meta-analysis of 15 studies comprising 35,551 participants with varied ethnicities to determine the optimum 1-h PG level equivalent to the 2-h PG ≥11.1 mmol/L diagnostic of type 2 diabetes.

RESEARCH DESIGN AND METHODS

Search Strategy and Selection Criteria

In this meta-analysis, principal investigators of 15 studies involved in cross-sectional or longitudinal studies from 1965 to the present with access to FPG, 1-h PG (index test), and 2-h PG (reference standard) data during an OGTT participated (10,15–24). Cases of diabetes included newly detected type 2 diabetes defined as 2-h PG ≥11.1 mmol/L during the OGTT. We excluded participants identified to have diabetes only based on FPG ≥7.0 mmol/L and/or HbA_{1c} ≥48.0 mmol/mol (6.5% [≥43.0 mmol/mol (6.1%) for Japanese participants]) or who were on glucose-lowering medications (25). This is because we considered 2-h PG as the reference standard and including participants based on FPG and/or HbA_{1c} criteria would have reduced the specificity of the 1-h PG to detect diabetes with a 2-h PG ≥11.1 mmol/L. Individuals without diabetes from the same cohorts (2-h PG <11.1 mmol/L) constituted the control group.

Data Analysis

In 8 of the 15 studies included in the meta-analysis, analysts provided information on the study design, sample size, setting (e.g., primary health centers, diabetes clinics, population based), recruitment procedures, percentage of women, mean age, mean 1-h PG, and percentage of diabetes cases based on the 2-h PG. Furthermore, they provided the numbers of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) and the 1-h PG cutoff for detection of diabetes based on a 2-h PG ≥11.1 mmol/L (Supplementary Material). We included 1-h PG thresholds at the maximum Youden index and the minimum distance for each study, if they differed. The primary analyst (V.A.) performed the analyses using raw data in seven studies. Two authors (V.A. and M.B.) independently performed a quality assessment of the studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2 tool through consensus (26) (Supplementary Material).

We constructed a forest plot displaying TP, FP, FN, TN, 1-h PG cutoffs, sensitivity, specificity, and their 95% CI of 1-h PG for each study using Review Manager (RevMan) 5.3. We further created a receiver operating characteristic (ROC) ellipse plot that depicts the estimate of each study with its 95% confidence region (CR) in the ROC area. Furthermore, we constructed a forest plot of 1-h PG showing the log diagnostic odds ratio (lnDOR) of each study with its summary estimate. In addition, we plotted a Fagan nomogram that integrates prevalence, likelihood ratios (positive and negative likelihood ratio), and posttest probabilities (positive [PPV] and negative predictive values).

Having two cutoffs from each study and two outcomes as specificities and sensitivities makes this a multilevel random effects model. To account for this structure, while meta-analyzing, we constructed a summary ROC curve for 1-h PG using a class of weighted mixed linear effects regression model that modeled sensitivities and specificities separately for participants with and without diabetes across all studies considering fixed effects for studies, cutoffs, and their interactions and various random effects (27). Furthermore, we used three different λ while constructing the summary ROC curve in order to assign different weights to specificities and sensitivities: 1/2, weighing specificity and sensitivity equally and thus resembling the

[16]German Center for Diabetes Research (DZD), Neuherberg, Germany
[17]Department of Clinical and Molecular Medicine, Sapienza University of Rome, Rome, Italy
[18]Department of Internal Medicine, Hokuriku Central Hospital, Toyama, Japan
[19]Lund University Diabetes Centre, Lund University, Malmö, Sweden
[20]Public Health Promotion Unit, Finnish Institute for Health and Welfare, Helsinki, Finland
[21]Department of Public Health, University of Helsinki, Helsinki, Finland
[22]Saudi Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi Arabia
[23]Department of Public Health, Clinicum, University of Helsinki, Helsinki, Finland
[24]Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, MA
[25]Division of Endocrinology and Metabolism, Department of Medicine and Department of Population Health, and NYU Langone Diabetes Prevention Program, NYU Grossman School of Medicine, New York, NY
[26]Abdominal Centre, Endocrinology, Helsinki University Hospital, and Folkhalsan Research Centre, Biomedicum, and Research Program Unit, Clinical and Molecular Medicine, University of Helsinki, Helsinki, Finland

Corresponding author: Vasudha Ahuja, vasudha.ahuja@helsinki.fi

Received 7 July 2020 and accepted 11 January 2021

This article contains supplementary material online at https://doi.org/10.2337/figshare.13570910.

M.B. and T.T. contributed equally.

maximum Youden index; 2/3, enhancing sensitivity; and 1/3, emphasizing specificity. In addition, among models with the same fixed and different random effects, we chose the model that best described our data using the smallest restricted maximum likelihood criteria. Finally, to find the optimum cutoff of 1-h PG for detection of 2-h PG ≥11.1 mmol/L on the chosen random effect model, we chose the λ that provided the best combination of specificity and sensitivity.

We explored heterogeneity across the studies using meta-regression with the design of studies (cross-sectional vs. longitudinal), the setting of studies (diabetes clinic vs. population based), the dose of glucose used for the OGTT (75 g vs. 100 g), ethnicity (South Asian, American Indian, Japanese, and Mexican American vs. Caucasian), and bias (studies with risk vs. low risk of bias) as covariates. We also examined sample size–related effects by constructing a funnel plot and assessed its asymmetry using the test of Deeks et al. (28).

Finally, for the seven studies with available raw data, we performed certain subanalyses. First, we restricted the analyses to the cases of diabetes with 2-h PG just above the diagnostic cutoff (≥11.1 to ≤13.0 mmol/L) not only to increase the specificity of the analysis for that cutoff but also because they are likely to be of more recent onset than those with the 2-h PG ≥11.1 mmol/L. Second, we compared the cutoff obtained by meta-analysis of unadjusted cutoffs with the cutoff obtained by meta-analysis of age-, sex-, and BMI-adjusted cutoffs in order to assess how these factors affect the cutoff of the 1-h PG. We used R, version 3.6.3, for analyses unless mentioned otherwise.

RESULTS

We included 15 studies with 35,551 participants representing Caucasian, American Indian, Japanese, Mexican American, and South Asian ethnicities (46.2% non-Caucasian). Four studies were longitudinal, and 11 were cross-sectional; two were primary health care center based, four diabetes clinic based, and nine population based. All but one study used a glucose dose of 75 g for the OGTT (Table 1). The mean value of 1-h PG across studies ranged from 10.1 to 18.5 mmol/L in individuals with and 7.4–9.2 mmol/L in those without diabetes. Of the newly detected cases of diabetes (N = 3,382), we excluded 677 (20.0%) who had

Table 1—Characteristics of included studies

Study*	Design	Setting	Ethnicity	*N* (% females)	Age at baseline (years)	Glucose dose in OGTT (g)	1-h PG mmol/L	Type 2 diabetes cases, 2-h PG ≥11.1 mmol/L
BFS, 1990 (15)†‡	Cross-sectional	Primary health care	Caucasian	2,995 (55)	46.2 ± 13.7	75	7.9 ± 2.7	126 (4.2)
BPS, 1990 (15)†‡	Longitudinal	Primary health care	Caucasian	3,168 (55)	54.0 ± 14.7	75	8.0 ± 2.7	85 (2.7)
CATAMERI, 2005 (16)	Cross-sectional	Diabetes clinic	Caucasian	3,324 (54)	48.4 ± 13.9	75	8.8 ± 2.7	249 (7.5)
DIAGEN, 1996 (17)‡	Cross-sectional	Population	Caucasian	2,679 (56)	52.6 ± 16.5	75	9.2 ± 2.9	204 (7.6)
DIAPASON, 2014 (18)‡	Cross-sectional	Diabetes clinic	Caucasian	531 (57)	59.4 ± 9.9	75	8.4 ± 2.6	34 (6.4)
GENFIEV, 2003 (19)	Cross-sectional	Diabetes clinic	Caucasian	916 (57)	49.3 ± 11.3	75	9.8 ± 2.8	116 (16.6)
GOH, 1979 (10)	Cross-sectional	Population	Caucasian	2,092 (48)	51.3 ± 8.0	100	8.6 ± 3.2	149 (7.1)
HPS, 1966 (20)†	Cross-sectional	Population	Caucasian	1,026 (0)	44.0 ± 7.7	75	7.1 ± 2.0	11 (0.9)
MDRF, 1991 (21)	Cross-sectional	Diabetes clinic	South Asian	9,651 (45)	45.0 ± 12.0	75	9.4 ± 2.5	802 (8.3)
Oulu45, 2001 (22)†	Cross-sectional	Population	Caucasian	933 (56)	56.8 ± 0.6	75	8.6 ± 2.3	33 (3.6)
Oulu45P, 2001 (22)†	Longitudinal	Population	Caucasian	825 (58)	56.8 ± 0.6	75	8.0 ± 1.9	44 (5.3)
PIBS, 1966 (14)	Longitudinal	Population	American Indian	2,664 (50)	32.2 ± 15.1	75	8.2 ± 4.1	399 (15.1)
PSW, 2006 (23)‡	Cross-sectional	Population	Japanese	2,085 (32)	52.6 ± 7.2	75	8.5 ± 2.6	70 (3.4)
PSWP, 2006 (23)‡	Longitudinal	Population	Japanese	1,997 (28)	52.4 ± 6.9	75	8.6 ± 2.7	65 (3.23)
SAHS, 1992 (24)	Cross-sectional	Population	Mexican American	689 (66)	49.8 ± 12.1	75	12.1 ± 4.2	318 (46.2)

Data are means ± SD or *n* (%) unless otherwise indicated. BFS, Botnia Family Study; BPS, Botnia Prospective Study; CATAMERI, CATAnzaro MEtabolic RIsk factors; DIAGEN, DIAbetes GENetic study; DIAPASON, Diabetes Prediction and Screening Observational; GENFIEV, GENetics, pathoPHYsiology, and EVolution of type 2 diabetes; GOH, Israel Study of Glucose Intolerance, Obesity and Hypertension; HPS, Helsinki Policemen Study; MDRF, Madras Diabetes Research Foundation; Oulu45P, Oulu45 Prospective; PIBS, Pima Indian Biennial Study; PSW, Public School Worker; PSWP, Public School Worker Prospective; SAHS, San Antonio Heart Study. *Studies with initiation year. †Blood glucose converted to plasma glucose with a conversion factor of 1.13. ‡Studies with data for HbA_{1c}.

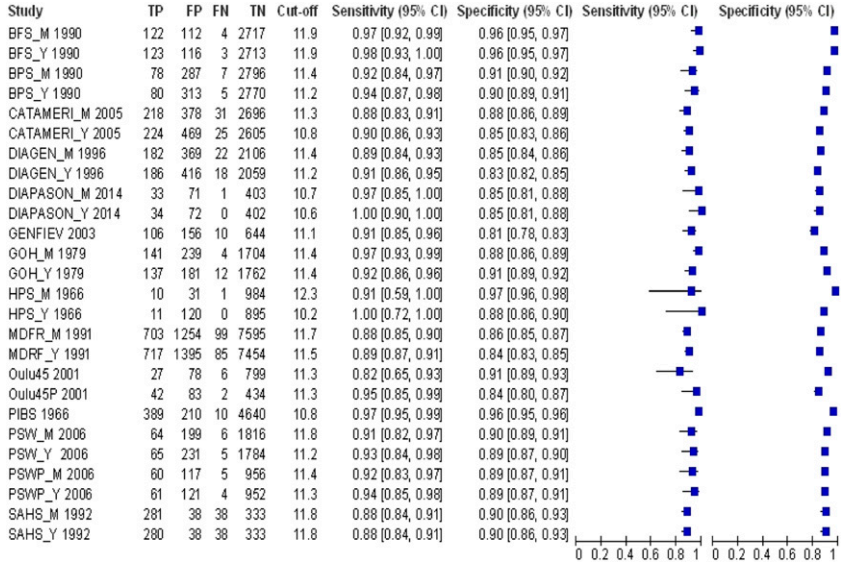

Figure 1—A forest plot showing the sensitivity and specificity of the obtained 1-h PG cutoffs to detect diabetes (defined as 2-h glucose ≥11.1 mmol/L) in the individual studies together with the number of participants with TP, FP, FN, and TN results. _M after the study name indicates the cutoff at the minimum distance and _Y at the Youden index (in case of no postfix, the cutoff is the same at the minimum distance and Youden index). BFS, Botnia Family Study; BPS, Botnia Prospective Study; CATAMERI, CATAnzaro MEtabolic RIsk factors; DIAGEN, DIAbetes GENetic study; DIAPASON, Diabetes Prediction and Screening Observational; GENFIEV, GENetics, pathoPHYsiology, and EVolution of type 2 diabetes; GOH, Israel Study of Glucose Intolerance, Obesity and Hypertension; HPS, Helsinki Policemen Study; MDRF, Madras Diabetes Research Foundation; Oulu45P, Oulu45 Prospective; PIBS, Pima Indian Biennial Study; PSW, Public School Worker; PSWP, Public School Worker Prospective; SAHS, San Antonio Heart Study.

diabetes based on FPG and/or HbA_{1c} only and analyzed data for 2,705 (80.4%) of diabetes based on a 2-h PG ≥11.1 mmol/L: 1,746 (51.6%) of these based on 2-h PG only and 959 (28.4%) based on both 2-h PG and FPG/HbA_{1c} (Supplementary Table 1).

QUADAS-2 assessment showed a strong quality of evidence (Supplementary Fig. 1). Eleven studies had low risk of bias or applicability concerns, while two studies were at risk for bias in the domain of patient selection and two in applicability concerns.

The forest plot shows that 1-h PG of 10.2–11.9 mmol/L had a sensitivity of 0.82–1.0 and specificity of 0.79–0.97 to detect a 2-h PG ≥11.1 mmol/L (Fig. 1). In the ROC ellipse plot, the estimates from all studies positioned in the upper-left portion of the ROC area demonstrate high diagnostic accuracy of the 1-h PG in detection of 2-h PG ≥11.1 mmol/L (Fig. 2). The forest plot of lnDOR shows 4.6 times higher odds of obtaining a positive result with use of 1-h PG in individuals with than without diabetes (Supplementary Fig. 2). The Fagan nomogram indicates that the probability of having diabetes increases from 7% to 45% with a positive result. Furthermore, the probability of having diabetes drops from 7% to 1% with a negative result (Supplementary Fig. 3*A*).

The selected different random slope model suggested a study-specific effect of 1-h PG on the accuracy in detection of 2-h PG ≥11.1 mmol/L. Supplementary Fig. 4 shows three alternative cutoffs at different λ levels: 10.6 mmol/L (95% CI 10.0, 11.3) at λ 2/3 (higher sensitivity), 11.6 mmol/L (10.6, 12.6) at λ 1/2 with equal weights for sensitivity and specificity (Youden index), and 12.5 mmol/L (11.3, 14.0) at λ 1/3 (higher specificity). At these cutoffs (10.6 vs. 11.6 vs. 12.5 mmol/L) the 1-h PG had sensitivity of 0.95 (0.91, 0.97) vs. 0.92 (0.87, 0.95) vs. 0.87 (0.79, 92) and specificity of 0.86 (0.82, 0.89) vs. 0.91 (0.88, 0.93) vs. 0.94 (0.92, 0.96), respectively. At all these cutoffs, the area under the curve of 1-h PG in detection of 2-h PG ≥11.1 mmol/L was 0.939 (95% CR for sensitivity at given specificity 0.904, 0.946). Table 2 shows the numbers of TP, FN, FP, TN, and PPV at these cutoffs. As expected, the number of FN increased, FP decreased, and PPV increased as cutoff levels of 1-h PG increased.

Cross-sectional studies are more likely to recruit long-standing undiagnosed cases of diabetes as "incident" compared with longitudinal studies, and clinic-based studies are likely to recruit more cases of IH compared with population-based studies. However, the meta-regression analysis did not show differences in sensitivity or specificity for the diagnostic accuracy of the 1-h PG to diagnose the 2-h PG ≥11.1 mmol/L in comparison of cross-sectional and longitudinal studies (P = 0.43 and 0.88, respectively) or diabetes clinic-based and population-based studies (P = 0.58 and 0.46, respectively). Further, studies with administration of 75 g glucose demonstrated no difference in diagnostic accuracy compared with the study with a 100-g dose (sensitivity P = 0.88, specificity P = 0.24). In addition, meta-regression analysis by ethnicity showed that American Indians had the highest sensitivity of 1-h PG

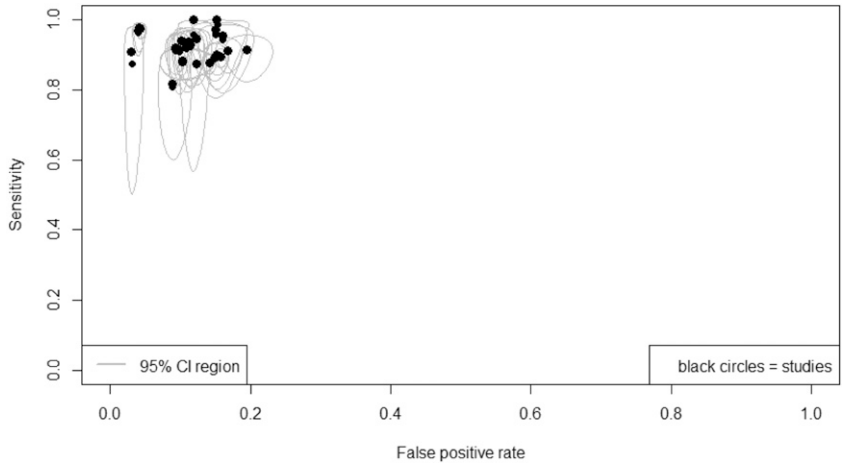

Figure 2—ROC ellipse plot showing the cutoffs of studies with 95% CRs.

Table 2—The number of TP and FP or TN and FN cases with three different 1-h PG cutoffs in diagnosis of type 2 diabetes of 2-h PG ≥11.1 mmol/L and the associated PPV

Cutoff in mmol/L (Se, Sp)	Weight ratio for Se vs. Sp	λ	Case subjects, type 2 diabetes by 2-h PG (N = 2,705)		Control subjects, by 2-h PG (N = 34,246)		PPV % FP/TP + FP
			TP	FN	FP	TN	
10.6 (0.95, 0.86)	More	2/3	2,570	135	4,794	29,452	34.9
11.6 (0.92, 0.91)	Equal	1/2	2,489	216	3,082	31,164	44.6
12.5 (0.87, 0.94)	Less	1/3	2,353	352	2055	32,191	53.4

Se, sensitivity; Sp, specificity.

followed by Japanese, Caucasians, South Asians, and Mexican Americans, respectively ($P < 0.0001$). Again, American Indians had the highest specificity of 1-h PG, followed by Caucasians, Mexican Americans, Japanese, and South Asians, respectively ($P < 0.0001$). Although studies with risk of bias demonstrated sensitivity similar to that of studies with low risk of bias ($P = 0.19$), they had lower specificity ($P = 0.001$) (Supplementary Tables 2 and 3). Furthermore, the examination of the funnel plot with use of the Deeks test showed nonsignificant ($P = 0.21$) results indicating the absence of sample size–related effects (Supplementary Fig. 5).

Using the raw data of seven studies in the subanalysis restricted to detect diabetes with 2-h PG ≥11.1 mmol/L to ≤13.0 mmol/L, presumed to be of fairly recent duration, we found that the cutoff levels of the 1-h PG were higher than when individuals with 2-h PG ≥11.1 mmol/L were included: 12.6 mmol/L at λ 2/3 (higher sensitivity), 13.5 mmol/L at λ 1/2 (Youden index), and 14.5 mmol/L at λ 1/3 (higher specificity). Furthermore, the sensitivity of the 1-h PG to detect diabetes with 2-h PG within ≥11.1 to ≤13.0 mmol/L was lower and specificity was higher than for detection of diabetes with the 2-h PG ≥11.1 mmol/L (Supplementary Table 4). Finally, we found that unadjusted cutoffs were either similar to or lower than age-, sex-, and BMI-adjusted cutoffs depending on the λ used (Supplementary Table 5).

CONCLUSIONS

In this meta-analysis of >35,000 individuals across multiple ethnic groups, we demonstrate that the 1-h PG of 10.6–12.5 mmol/L detects individuals with a 2-h PG level diagnostic of diabetes (≥11.1 mmol/L) with 87–95% sensitivity and 86–94% specificity. The choice of the 1-h PG cutoff depends on whether more weight is given to sensitivity or specificity. Thus, with the cutoff at the Youden index (11.6 mmol/L) with sensitivity of 92% and specificity of 91%, the 1-h PG detected 2,489 of 2,705 (92%) cases of type 2 diabetes while missing 216 (8%). Whereas the 1-h PG correctly classified 31,164 of 32,246 (91%) individuals as not having diabetes, it classified as many as 3,082 (9%) individuals who do not have diabetes by current criteria as having diabetes.

The OGTT is considered the "gold standard" for the diagnosis of diabetes despite having a large coefficient of variation and being inconvenient (29). It is noteworthy that it reflects the progressive failure of β-cell function, the primary phenomenon that drives the development of overt diabetes (30). While the clinical use of OGTT usually includes only FPG and 2-h PG levels, the deterioration of the insulin secretory response can be estimated from glucose and insulin concentrations either at 30 min postchallenge, as a proxy for first-phase insulin response, or at 2 h, reflecting both first-and second-phase insulin responses (31). Expectedly, 1-h PG, not currently measured during the OGTT, has a stronger correlation with the Matsuda index, the disposition index at 120 min, and glucose area under the curve than the 2-h PG (24). Considering that these proxy measurements of insulin secretion and insulin sensitivity are consistently lower in those who progress to type 2 diabetes, it is not surprising that the 1-h PG is also a more accurate predictor of progression to type 2 diabetes than IFG, IGT, and elevated HbA_{1c} (24,32).

In evaluation of the 1-h PG for diagnosing diabetes, two approaches can be considered. We evaluated the 1-h PG level coincident with the 2-h PG diagnostic of diabetes (11.1 mmol/L). The alternative and perhaps more biologically relevant approach would be to compare the 1-h PG with the 2-h PG value that best predicts diabetes-related complications. Regarding the approach described herein, the 1-h PG is convenient and strongly correlates with the 2-h PG. However, several factors affect the relationship between the 1-h and 2-h PG. Different pathogenic mechanisms in glucose responsiveness and insulin secretion manifest in a significantly different ratio of 1-h PG and 2-h PG in carriers of *GCK* and *HNF1A* mutations (33). Furthermore, the profile of the glucose response during the OGTT changes with progression from normoglycemia to IGT to overt diabetes (30). Finally, glucose control may have an effect, since chronic hyperglycemia causes an insulin secretory defect that can result in different cutoff values in cohorts with recent-onset or long-standing diabetes (34). For this reason, in this study, we only included newly detected cases based on screening with OGTT and a 2-h PG value diagnostic of diabetes. This may partly explain the lower cutoff for the 1-h PG of 11.6 mmol/L in the present meta-analysis compared with 13.0 mmol/L in a Chinese hospital-based study (13).

The current diagnostic threshold values for diabetes originated based on the association of glycemic levels with increased prevalence of diabetic retinopathy, especially nonproliferative diabetic retinopathy (35). In this regard, the 1-h PG was significantly associated with prevalent and incident diabetic retinopathy in American Indians and with incident diabetic retinopathy in a Swedish cohort

(8,14). Furthermore, the 1-h PG was associated with diabetic retinopathy similarly to association of 2-h PG with diabetic retinopathy in the former population. Multiple studies have demonstrated an association of 1-h PG with cardiovascular outcomes and mortality (8–10). Moreover, among men without diabetes in the Malmö Preventive Project, the 1-h PG predicted cardiovascular death and all-cause mortality better than the 2-h PG (8).

As information relating to the presence of retinopathy or cardiovascular disease was not available for the cohorts analyzed in this meta-analysis, we could not evaluate the cutoff of 1-h PG that would best detect diabetes complications. However, Paddock et al. (14), in a cross-sectional analysis, identified a 1-h PG threshold of 12.0 mmol/L for diagnosing type 2 diabetes in American Indians with retinopathy, comparable with 11.6 mmol/L in the present meta-analysis. In the same study, the cutoff based on a longitudinal analysis was 12.8 mmol/L, again similar to the 12.6 mmol/L at λ 2/3 in our study when we restricted the analysis to presumably more recently diagnosed type 2 diabetes (14). Ideally, the comparisons should include information regarding the distribution of values, as the results will differ substantially if the majority of individuals have 2-h PG near the cutoff of 11.1 mmol/L or much higher.

A meta-analysis only enables use of aggregate measures, e.g., proportion of females, therefore, assessing differences in diagnostic accuracy according to participant-level variables may introduce bias. Nevertheless, using raw data from the available seven studies, we explored how these demographic factors affect the diagnostic accuracy of 1-h PG. First, we found that the unadjusted and age-, sex-, and BMI-adjusted cutoffs of 1-h PG were significantly different in five out of seven studies (Supplementary Table 6). Second, at the meta-analytical level, we found the meta-analyzed unadjusted estimates to be either similar to or lower than meta-analyzed adjusted cutoffs (Supplementary Table 5). Additionally, the cutoffs differed minimally according to ethnicities, as did their sensitivities and specificities, except for American Indian, where the cutoff was lower and sensitivity and specificity were higher than in other groups, which may be due to exceptionally high risk of type 2 diabetes in this population (36). Moreover, universal diagnostic cutoff values for diagnosis of diabetes apply for all glycemic indices irrespective of age, sex, BMI, and ethnicity. This is true despite reported distinctive values of these indices in individuals of different demographic characteristics without diabetes (37). Thus, in line with the current universal diagnostic threshold values, we suggest using the same cutoff value of 1-h PG to detect type 2 diabetes among different groups.

The very first criteria for usefulness of a diagnostic test is its ability to discriminate between individuals with and without disease, i.e., the sensitivity and specificity. These are adequately high for the 1-h PG of 11.6 mmol/L with use of the 2-h PG for defining disease status. However, 55% of individuals classified as having diabetes by this 1-h PG did not have diabetes according to the 2-h PG (Table 2). Although the sensitivity and specificity are not mathematically dependent on prevalence, the number of FN increases and that of FP decreases as prevalence increases. Consequently, using the Pima Indian Biennial Study with a high prevalence of diabetes (15.1%) instead of all cohorts of the meta-analysis combined, with a lower prevalence (7%), increased the PPV from 45% to 64% (and decreased the number of FP from 55% to 36%) (19) (Supplementary Fig. 3*B*).

Some factors favor using a higher cutoff of 12.5 mmol/L. First, especially in populations with a low prevalence, a higher cutoff would be needed to increase the PPV; using 12.5 mmol/L instead of 11.6 mmol/L in the meta-analysis increased the PPV from 45% to 53% (Table 2). Second, in the subanalysis of individuals presumably having more recent-onset diabetes (2-h PG $<$13 mmol/L), the cutoff of 1-h PG was higher: $>$12.5 mmol/L (Supplementary Table 4). On the other hand, a large proportion of the individuals with a FP had 2-h PG values just below the current diagnostic cutoff for diabetes, and the majority (59%) had IGT (Supplementary Fig. 6). Previous studies reported that ~8% of people with IGT in the U.S. Diabetes Prevention Program (DPP) had evidence of diabetic retinopathy, suggesting a significant FN rate for the FPG and 2-h PG criteria and, thereby, underestimation of detection of dysglycemia (38). Furthermore, the 1-h PG has a stronger association with cardiovascular outcome and all-cause mortality than the 2-h PG (8). Thus, we hypothesize that the so-called FP cases, i.e., diabetes based on the 1-h PG but not the 2-h PG, may actually turn out to be TP cases regarding high risk of complications; those in this category serve as a target group for prevention.

It needs to be stressed that we are not proposing that the OGTT be performed as the initial screening test for type 2 diabetes (or prediabetes), as this would be highly infeasible and costly. In accordance with others, we advocate implementation of validated diabetes risk screening calculators (e.g., Finnish Diabetes Risk Score, ADA) to identify individuals at high risk (39). Further laboratory measurements would only be instituted in those identified as high risk based on the outcome of the screening calculator. The diagnosis of diabetes would be confirmed with a second test as recommended by WHO and ADA (40). With this procedure, the proportion of FP cases would likely be reduced that otherwise might incorrectly have had a diabetes diagnosis suggested. Furthermore, individuals who have been positively screened may have ongoing abnormalities in glucose regulation and therefore still remain at high risk for developing diabetes in the future and may benefit from lifestyle modification.

The strength of this meta-analysis is in its size (~35,000 participants) and diversity including populations from different countries. As we obtained raw estimates from studies in contrast to extracting published data, we achieved uniformity in defining type 2 diabetes and obtained complete information to assess the quality of studies. Two of the studies reported herein may have had volunteer bias due to convenient sampling, and two studies had significant loss to follow-up, which would have resulted in increased proportions of cases and decreased specificity in these studies compared with others. Overall, the quality of evidence was strong. The meta-analysis also has certain weaknesses. The number of studies included is small. Although we included studies having participants with different ethnic backgrounds, major ethnic groups such as of African or South American origin were missing. Moreover, it is ideal to choose diagnostic thresholds using incident cases of diabetes from population-based

longitudinal studies, as differences in the characteristics of participants in non-population-based and cross-sectional studies might affect the accuracy of a test. While we did not find significant differences in accuracy of the 1-h PG in detection of type between longitudinal and cross-sectional studies or between population and non-population-based studies, a higher 1-h PG cutoff was obtained in the subgroup with presumably more recent onset (2-h PG $<$13 mmol/L). Although examination of the funnel plot showed nonsignificant results, it displays an asymmetry that may point to a significant sample size–related effect. Here, a nonsignificant Deeks test might reflect its low power in case of heterogeneous DOR. Of note, the presence of sample size–related effects may reflect not only the possibility of publication bias (rather, in this meta-analysis it would reflect sampling bias) but also the relation of the sample size of the studies to the type of study population or study quality. However, after exclusion of the studies that stood apart in the Deeks funnel plot, the cutoff of 1-h PG was similar with little change in sensitivity and specificity (11.3 mmol/L [0.91, 0.89]; data not shown).

In summary, a 1-h PG of 11.6 mmol/L detected the 2-h PG $\geq$11.1 mmol/L diagnostic of type 2 diabetes with high sensitivity and specificity among adults previously undiagnosed with diabetes but detected a high proportion of FP cases. At least three things warrant further research including other ethnic backgrounds. First, we suggest reproducibility studies of 1-h PG compared with 2-h PG in populations other than American Indians, for whom reproducibility is poorer and the distribution bimodal in contrast to most populations (41). Second, we recommend population-based longitudinal studies comparing the strength of association of the 1-h PG and 2-h PG with diabetic retinopathy and other microvascular complications, cardiovascular complications, and all-cause mortality.

Funding and Duality of Interest. V.A. was funded through a grant from the Academy of Finland (grant 312072). The Botnia studies (L.G., T.T.) have been financially supported by grants from Folkhälsan Research Foundation, the Sigrid Juselius Foundation, the Academy of Finland (grants 263401, 267882, 312063 to L.G. and 312072 to T.T.), University of Helsinki, Nordic Center of Excellence in Disease Genetics, the European Union (EXGENESIS and MOSAIC projects, FP7-600914), Ollqvist Foundation, Swedish Cultural Foundation in Finland, Finnish Diabetes Research Foundation, Foundation for Life and Health in Finland, Signe and Ane Gyllenberg Foundation, Finnish Medical Society, Paavo Nurmi Foundation, State Research Funding via the Helsinki University Hospital, Perklén Foundation, Närpes Health Care Foundation, and Ahokas Foundation. The study has also been supported by the Ministry of Education in Finland, Municipal Health Care Center and Hospital in Jakobstad, and Health Care Centers in Vasa, Närpes, and Korsholm. The Botnia Study has received research funding from Pfizer Inc. and Regeneron Pharmaceuticals. The research leading to these results received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 269045. T.V.F. was supported by the European Foundation for the Study of Diabetes. DIAPASON (Diabetes Prediction and Screening Observational) was supported by Fondazione Romeo ad Enrica Invernizzi–Milano, an EFSD/Sanofi program (to L.L.S.), and Italian Ministry of Health–Ricerca Corrente. GENFIEV has been supported by FoRiSID, Rome, Italy, and an unconditional grant from Eli Lilly, Italy. GOH (Israel Study of Glucose Intolerance, Obesity and Hypertension) was funded by The DCURE Foundation. Oulu45 and Oulu45P studies were funded by the Finnish Cultural Foundation Central Fund. PIBS (Pima Indian Biennial Study) was funded by the Division of Intramural Research of National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health. The PSW (Public School Worker) and PSWP (Public School Worker Prospective) studies were funded by the Japan Health Promotion Foundation and Grant-in-Aid from Toyama Medical Association. A.Ce. receives research support from Mitsubishi; is a member of the advisory boards of Abbott, BD Biosciences, Eli Lilly, Janssen, and Mundipharma; and is a member of the speakers bureaus of AstraZeneca, Berlin Chemie, Boehringer Ingelheim, Novo Nordisk, and Roche Diagnostics. C.B. receives honoraria for consulting from Novo Nordisk and research support from Roche Diagnostics. J.T. received research grants from Bayer Pharma, Boehringer Ingelheim, Merck, and Sanofi; received consulting and travel fees from Eli Lilly, Merck, Merck Sharp & Dohme, and Novo Nordisk; and owns stocks of Orion Pharma. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. V.A., M.B., and T.T. designed the meta-analysis. A.Ch., R.A.D., A.Ce., S.D.P., M.A.-G., S.K.-K., R.D., P.S., G.S., R.O., V.M., L.G., and T.T. designed the individual studies. T.A.P., A.Ch., L.L.S., R.M.A., S.J., T.V.F., P.T., R.A.D., A.Ce., S.D.P., M.A.-G., R.D., P.H.B., W.C.K., P.S., G.S., and R.O. contributed to data collection. V.A. and M.B. performed the risk-of-bias analysis. V.A., P.A., H.L., A.Ch., A.H.B., A.J., C.B., L.L.S., R.P., U.V., V.B., and R.D. performed the analyses of the individual studies. V.A. and P.A. performed the meta-analysis supervised by S.R. and T.T. V.A., P.A., M.B., and T.T. wrote the manuscript. All authors contributed to data interpretation and revised the report. V.A. and T.T. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the 55th Annual Meeting of the European Association for the Study of Diabetes, 16–20 September 2019, Barcelona, Spain, and at the 80th Scientific Sessions of the American Diabetes Association, 12–16 June 2020.

References

1. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. Diabetes 1979; 28:1039–1057
2. World Health Organization. *WHO Expert Committee on Diabetes Mellitus [Meeting Held in Geneva From 25 September to 1 October 1979]: Second Report*. Geneva,World Health Org., 1980
3. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 1997;20:1183–1197
4. World Health Organization. *Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications: Report of a WHO Consultation. Part 1, Diagnosis and Classification of Diabetes Mellitus*. Geneva,World Health Org., 1999
5. Gillett MJ. International Expert Committee report on the role of the A1c assay in the diagnosis of diabetes: Diabetes Care 2009; 32(7): 1327-1334. Clin Biochem Rev 2009;30:197–200
6. World Health Organization. *Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus: Abbreviated Report of WHO Consulatation*. Geneva, World Health Org., 2011
7. Bergman M, Manco M, Sesti G, et al. Petition to replace current OGTT criteria for diagnosing prediabetes with the 1-hour post-load plasma glucose $\geq$ 155 mg/dl (8.6 mmol/L). Diabetes Res Clin Pract 2018;146:18–33
8. Pareek M, Bhatt DL, Nielsen ML, et al. Enhanced predictive capability of a 1-hour oral glucose tolerance test: a prospective population-based cohort study. Diabetes Care 2018;41:171–177
9. Lind M, Tuomilehto J, Uusitupa M, et al. The association between HbA1c, fasting glucose, 1-hour glucose and 2-hour glucose during an oral glucose tolerance test and cardiovascular disease in individuals with elevated risk for diabetes. PLoS One 2014;9:e109506
10. Bergman M, Chetrit A, Roth J, Dankner R. One-hour post-load plasma glucose level during the OGTT predicts mortality: observations from the Israel Study of Glucose Intolerance, Obesity and Hypertension. Diabet Med 2016;33:1060–1066
11. Glucose tolerance and mortality: comparison of WHO and American Diabetes Association diagnostic criteria. The DECODE study group. European Diabetes Epidemiology Group. Diabetes Epidemiology: Collaborative analysis Of Diagnostic criteria in Europe. Lancet 1999;354:617–621
12. Meigs JB, Nathan DM, D'Agostino RB Sr., Wilson PW; Framingham Offspring Study. Fasting and postchallenge glycemia and cardiovascular disease risk: the Framingham Offspring Study. Diabetes Care 2002;25:1845–1850
13. Zhou W, Gu Y, Li H, Luo M. Assessing 1-h plasma glucose and shape of the glucose curve during oral glucose tolerance test. Eur J Endocrinol 2006;155:191–197

14. Paddock E, Looker HC, Piaggi P, Knowler WC, Krakoff J, Chang DC. One-hour plasma glucose compared with two-hour plasma glucose in relation to diabetic retinopathy in American Indians. Diabetes Care 2018;41:1212–1217
15. Tripathy D, Carlsson M, Almgren P, et al. Insulin secretion and insulin sensitivity in relation to glucose tolerance: lessons from the Botnia Study. Diabetes 2000;49:975–980
16. Succurro E, Arturi F, Caruso V, et al. Low insulin-like growth factor-1 levels are associated with anaemia in adult non-diabetic subjects. Thromb Haemost 2011;105:365–370
17. Abdul-Ghani MA, Abdul-Ghani T, Müller G, et al. Role of glycated hemoglobin in the prediction of future risk of T2DM. J Clin Endocrinol Metab 2011;96:2596–2600
18. La Sala L, Mrakic-Sposta S, Tagliabue E, et al. Circulating microRNA-21 is an early predictor of ROS-mediated damage in subjects with high risk of developing diabetes and in drug-naïve T2D. Cardiovasc Diabetol 2019;18:18
19. Bianchi C, Miccoli R, Bonadonna RC, et al.; GENFIEV Investigators. Pathogenetic mechanisms and cardiovascular risk: differences between HbA_{1c} and oral glucose tolerance test for the diagnosis of glucose tolerance. Diabetes Care 2012;35:2607–2612
20. Pyörälä M, Miettinen H, Laakso M, Pyörälä K. Hyperinsulinemia predicts coronary heart disease risk in healthy middle-aged men: the 22-year follow-up results of the Helsinki Policemen Study. Circulation 1998;98:398–404
21. Sai Prasanna N, Amutha A, Pramodkumar TA, et al. The 1h post glucose value best predicts future dysglycemia among normal glucose tolerance subjects. J Diabetes Complications 2017; 31:1592–1596
22. Mutt SJ, Jokelainen J, Sebert S, et al. Vitamin D status and components of metabolic syndrome in older subjects from Northern Finland (latitude 65°North). Nutrients 2019;11:1229
23. Oka R, Yagi K, Sakurai M, et al. Insulin secretion and insulin sensitivity on the oral glucose tolerance test (OGTT) in middle-aged Japanese. Endocr J 2012;59:55–64
24. Abdul-Ghani MA, Williams K, DeFronzo RA, Stern M. What is the best predictor of future type 2 diabetes? Diabetes Care 2007;30:1544–1548
25. Kashiwagi A, Kadowaki T, Haneda M, et al. Consensus and statement on international standardization of HbA1C in Japan: committee report on diabetes mellitus laboratory testing standardization. J Jpn Diabetes Soc 2009;52:811–818
26. Whiting PF, Rutjes AW, Westwood ME, et al.; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–536
27. Steinhauser S, Schumacher M, Rücker G. Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. BMC Med Res Methodol 2016;16:97
28. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 2005;58:882–893
29. World Health Orgnaization, International Diabetes Federation. *Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycaemia: Report of a WHO/IDF Consultation*. Geneva, World Health Org., 2006
30. Bonadonna RC, Boselli L, Dei Cas A, Trombetta M. Methods to assess in vivo insulin sensitivity and insulin secretion. In *Diabetes Epidemiology, Genetics, Pathogenesis, Diagnosis, Prevention, and Treatment (Endocrinology)*. Bonora E, DeFronzo RA, Eds. Switzerland, Springer, 2018
31. Kettunen JLT, Tuomi T. Human physiology of genetic defects causing beta-cell dysfunction. J Mol Biol 2020;432:1579–1598
32. Peddinti G, Bergman M, Tuomi T, Groop L. 1-hour post-OGTT glucose improves the early prediction of type 2 diabetes by clinical and metabolic markers. J Clin Endocrinol Metab 2019; 104:1131–1140
33. Stride A, Vaxillaire M, Tuomi T, et al. The genetic abnormality in the beta cell determines the response to an oral glucose load. Diabetologia 2002;45:427–435
34. White MG, Shaw JA, Taylor R. Type 2 diabetes: the pathologic basis of reversible β-cell dysfunction. Diabetes Care 2016;39:2080–2088
35. International Expert Committee. International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care 2009;32:1327–1334
36. Knowler WC, Pettitt DJ, Saad MF, Bennett PH. Diabetes mellitus in the Pima Indians: incidence, risk factors and pathogenesis. Diabetes Metab Rev 1990;6:1–27
37. Cavagnolli G, Pimentel AL, Freitas PA, Gross JL, Camargo JL. Effect of ethnicity on HbA1c levels in individuals without diabetes: systematic review and meta-analysis. PLoS One 2017;12: e0171315
38. Diabetes Prevention Program Research Group. The prevalence of retinopathy in impaired glucose tolerance and recent-onset diabetes in the Diabetes Prevention Program. Diabet Med 2007;24:137–144
39. Nowak C, Ingelsson E, Fall T. Use of type 2 diabetes risk scores in clinical practice: a call for action. Lancet Diabetes Endocrinol 2015;3:166–167
40. Zhang Y, Hu G, Zhang L, Mayo R, Chen L. A novel testing model for opportunistic screening of pre-diabetes and diabetes among U.S. adults. PLoS One 2015;10:e0120382
41. Rushforth NB, Bennett PH, Steinberg AG, Miller M. Comparison of the value of the two- and one-hour glucose levels of the oral GTT in the diagnosis of diabetes in Pima Indians. Diabetes 1975;24:538–546